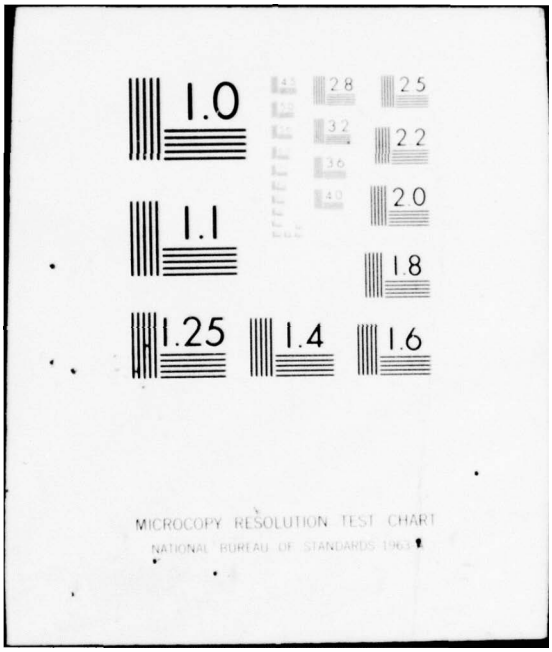
AD-A038 258 NORTHWESTERN UNIV EVANSTON ILL DEPT OF INDUSTRIAL E--ETC F/G 12/1
OPTIMAL MASS SCREENING UNDER CONSTANT AND THRESHOLD TEST RELIAB--ETC(U)
MAR 77 W P PIERSKALLA, J A VOELKER N00014-67-A-0356-0030
UNCLASSIFIED TR-5 NL

1 OF 1
AD A038258

END
DATE FILMED
5-77


1.0
1.1
1.25
1.4
1.6
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963

(12)

(6) OPTIMAL MASS SCREENING UNDER CONSTANT AND THRESHOLD TEST RELIABILITIES.

by

(10) William P. Pierskalla*
and
John A Voelker**

(9) Technical Report, No. 5

(11) March 1977

(12) 42p.

(14) TR-5

(15) Prepared under Office of Naval Research
Contracts N00014-67-A-0356-0030 (Task NR042-322)
and N00014-75-C-0797 (Task NR042-322)
W. P. Pierskalla, Project Director

*Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois 60201

**Corporate Research Department, Montgomery Ward Company, Chicago, Illinois 60680

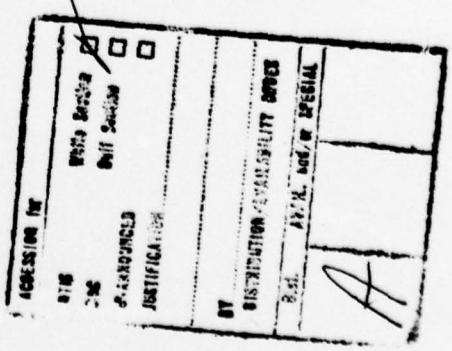
AB

ABSTRACT

This paper considers a population (composed, for example, of people, machines, or livestock) subject to a randomly occurring defect or disease. If there exist testing procedures capable of detecting the defect before it would otherwise become known, and if such early detection provides benefit, the periodic administration of such a test procedure to the members of the population, i.e., a mass screening program, may be advisable. This paper develops and analyzes two important cases of a general model of a mass screening program. In both cases, the defect arrival process is Poisson.

In the first case, the reliability of the test, i.e., the probability that the test will detect the disease, is dependent only on the type of test being administered and is constant over all values of elapsed time since the incidence of the defect. The long-run expected disutility per unit time is derived and variations of the disutility with regard to test reliability and testing frequency are presented. In addition, explicit solutions are provided for various special forms of the disutility function.

In the second case, the test will detect the defect if and only if the elapsed time since incidence exceeds a critical threshold T which characterizes the test. An expression for the long-run expected disutility per unit time is derived and decision rules to select between two alternative test types and costs are presented.



OPTIMAL MASS SCREENING UNDER CONSTANT AND THRESHOLD TEST RELIABILITIES

Mass screening is the process of inspecting a large population for defects. This population may consist of many subpopulations each with different numbers of members but more importantly each with different levels of susceptibility to the defects. If the early detection of a defect provides benefit, it may be advantageous to employ a test capable of revealing the defect's existence in its earlier stages. (Throughout this paper the words "defect" and "unit" or "individual" will refer to defect, disorder or disease and to a member of the population, respectively.)

Defects may arrive in a seemingly random fashion such as many types of machine failure, the incidence of certain types of cancer, diabetes, glaucoma, heart disease, etc., or they may arrive as the result of some contagion such as smallpox, polio, etc. It is the former type of arrival process, random arrivals, which is studied in this paper.

Now continuous monitoring would provide the most immediate detection of a defect. But considerations of expense and practicality will frequently rule out continuous monitoring so that a schedule of periodic testing--a screening program--may be the most practical means of achieving early detection of the defect. In general terms, the question then becomes one of how best to trade off the expense of testing which increases both with the frequency of test applications and with the cost of the type of test used against the benefits to be achieved from detecting the defect in an earlier stage of development.

The benefits of early detection depend upon the application considered. For example, in a human population being screened for some chronic disease the benefits of early detection might include an improved probability of ultimate cure, diminished time period of disability, discomfort, and loss of earnings, and reduced treatment cost. If the population being screened consists of machines engaged in some kind of production, the benefits of early detection might include a less costly ultimate repair and a reduction in the time period during which a faulty product is being unknowingly produced. If the population being screened consists of machines held in readiness to meet some emergency situation, an early detection of a defect would reduce the time the machine was not serving its protective function.

The expense of testing includes easily quantifiable economic costs such as those of the labor and materials needed to administer the testing. However, there can also be other important cost components which are more difficult to quantify. For example, in the case of a human population subject to medical screening, the cost of testing includes the inconvenience and possible discomfort necessitated by the test; the cost of false positives which entails both emotional distress and the need to do unnecessary follow-up testing; and even the risk of physical harm to the testee, e.g., from the cumulative effect of X-ray exposure.

In an earlier paper by Pierskalla and Voelker [1976], a mathematical model of a mass screening program over Q subpopulations was developed. The model allows for various testing technologies, with different reliability characteristics and costs, different arrival rates of the defect for each subpopulation, and budget, personnel and facility constraints.

The objective is to minimize the long-run disutility per unit time for the population. The decision variables are the testing frequencies for each subpopulation including the consideration whether a mass screening program is justified at all for any given subpopulation.

In this paper, this model is very briefly presented and then used to consider two important special cases of mass screening situations. The first case is that of a test or tests which have constant test reliability. This case is often one in which the testing instrument (machine, vaccine, etc.) has a relatively constant probability of false negatives or is one where the test may be perfectly reliable but the machine or device randomly produces defective units with a relatively constant probability.

The second special case of the general model is that of a test or tests which have a threshold reliability. In this case, there is an elapsed time $T \geq 0$ after the arrival of the defect such that the test is not able to detect the defect up to time T but will certainly detect the defect once the defect has matured beyond time T. This case is common in practice where there is a threshold of the defect before the test can recognize its existence. This threshold time T is, of course, a function of the test being administered.

Before proceeding to the cases of constant and threshold test reliabilities, it is useful to briefly discuss some of the related literature on mass screening models.

Some models of mass screening have used a multi-state approach. The Shwartz and Galliher [1975] model was concerned with mass screening for breast cancer. They utilized the Bross-Blumenson model for the progressive

evolution of the disease through a set of states specific to breast cancer. Detection of the disease carries a certain associated increase in life expectancy which is determined by the state of the disease at the time of detection. Moreover, both the probabilities of detecting the disease by means of a screening test and of detecting it in some other way (e.g., the woman noticing a lump) are a function of the disease state. Upon calibrating the parameters of their model to empirical data, Shwartz and Galliher utilize computer simulation to determine the payoff due to particular longitudinal schedules, i.e., due to particular sequences of testing intervals for a woman over her adult lifetime.

In Voelker [1976], multi-state models were considered in which there are countably many defect states through which the disorder evolves according to a semi-Markov process. (Zelen and Weiss [1965] have shown a semi-Markov process to provide a good fit to the course of development of one disease which they studied in detail.) The probability of a test detecting an existing defect is a function of the state of the defect at the time the test is applied. Certain defect states are regarded as "self-detecting". The reward for detecting a defect through the use of a screening test is a function of the state of the defect at the time of the detection and how long the defect has been in that state. The semi-Markov process of defect evolution is independent of the time of incidence of the corresponding defect which is a Poisson process.

The results include formulas for the long-run expected utility per unit time of a screening program, and several relationships involving the "steady-state" distribution of undetected defect states which will arise in response to that screening program.

Most works in, and related to, the area of modeling mass screening have not utilized a multi-state approach, but rather a time based approach. Such models usually focus upon either the detection delay or the lead time characteristics of a screening program. Detection delay is the time between the incidence of the defect and its detection (whether that detection is the result of a screening test or of the defect becoming self-evident). Lead time is the time difference between the time of detection via a screening test and the time detection would otherwise have occurred had not a screening program been in existence. These time based models sometimes do make reference to defect states, but the number of states is confined to two, the preclinical state and the clinical state. In the preclinical state, the defect exists but has not been discovered. In the clinical state, the defect has become self-evident--thus rendering a screening test superfluous. When the defect occurs, it is first in the preclinical state and then in the clinical state. Hence, both detection delay and lead time are determined by the time of detection relative to the sojourn interval of the preclinical state.

Some early papers on a time-based approach are Derman [1961], Roeloffs [1963, 1967], Barlow, Hunter and Proschan [1963], McCall [1969], and Zelen and Feinleib [1969]. In three later papers, Keller [1974], Kirch and Klein [1974] and Pierskalla and Voelker [1976], models are presented which assume perfect test reliability. In these papers, optimal mass screening schedules are presented. Keller used the calculus of variations to characterize the schedule. Kirch and Klein give an inspection schedule which will minimize expected detection delay subject to a constraint on the expected number of examinations an individual would incur over a lifetime. Pierskalla and

Voelker, as mentioned before, minimize the expected disutility per unit time for the population.

The most closely related work to the results in this paper are the works of Lincoln and Weiss [1964] and Prorock [1973, 1976]. Lincoln and Weiss studied the statistical characteristics of detection delay under the assumption that the times of examinations form a renewal process and that the probability of detecting the defect, $p(t)$, is a function of the defect's age, t. They derive open form equations, similar to renewal type equations, which relate the density functions for the following entities: the probability of detection at a test application ($p(t)$), the time until the defect becomes potentially detectable and from this time the forward recurrence time to the first test, the probability of the event that at a particular time a test occurs and all prior tests had failed to detect the defect, and the detection delay. For the two special cases where $p(t)$ is a constant and where $p(t)$ is exponential, the moments for the detection delay are derived in closed form. For uniform testing intervals (and general $p(\cdot)$), the distribution and moments of the detection delay are computed. They also studied the case of perfect test reliability.

Prorok has established properties for the lead time in the case of n successive screening tests conducted at uniform intervals. He also derives the proportion of preclinical cases detected, both at a given test and over all n tests. The test is assumed to have a constant reliability not necessarily equal to one. In addition, he derives an estimator for the mean lead time. This estimator utilizes a mix of both theoretical quantities and quantities which would be empirically available. In the case of a perfectly reliable test, his estimator requires only the latter kind of input.

In this paper, as in the Pierskalla and Voelker [1976] paper, the model does not seek to minimize the detection or inspection delay but rather the objective function is (initially) an arbitrary increasing disutility function. The reason for choosing a general function is that the disutility experienced upon the delayed detection of a defect may well vary in a highly nonlinear way with the length of the delay.

The disutility associated with a particular detection will depend only on the elapsed time since the defect's incidence. The notation $D(t)$ will express the disutility incurred if detection occurs t units of time after incidence. $D(\cdot)$ is assumed throughout this paper to be a nonnegative increasing function.

In those cases where $\sup_{s \geq 0} D(s) < \infty$, there is a natural relation between the notion of a utility function and $D(\cdot)$; namely,

$$U(t) = \sup_{s \geq 0} D(s) - D(t) .$$

$U(t)$ is a decreasing function expressing the disutility avoided by detecting a defect t units of time after its incidence.

In the next section, the general model is presented. Section 2 contains the analysis of the case when the probability of detection is a constant over all values of elapsed time since incidence. The long-run expected disutility per unit time is derived; its differential qualities (with respect to variations in the test reliability parameter and testing frequency) are exhibited; and explicit solutions are provided for various special forms of $D(\cdot)$. Also for special forms of $D(\cdot)$, decision rules are presented to select between two different kinds of tests which differ with respect to their reliabilities and cost per application.

In section 3, it is assumed that the test will detect the defect if and only if the elapsed time since incidence exceeds (or equals) a critical threshold T which characterizes the test. An expression for the long-run expected disutility per unit time is derived, and decision rules are developed to select between two alternative test types which differ with respect to their critical threshold T and their cost per application. Some examples for linear and quadratic disutility are given in the final section.

The proofs of all results are given in the Appendix.

1. A General Model for Time-Based Mass Screening

This model was developed in Pierskalla and Voelker [1976] and a relatively simple expression for the objective function was established. The model is now briefly presented (full details may be found in the above paper).

It is assumed that the jth subpopulation is of a fixed size, N_j, and that the defects arrive according to a stationary Poisson process with rate $N_j\lambda_j$ (N_j could be and usually is a large but finite number). It may be somewhat more realistic to set the defect arrival rate proportional to the number of defect-free units, rather than to the total number of units in the population. However, it is also assumed that no defect can remain undetected longer than T^* after its arrival, even without any screening tests being given. Hence, $T^*N_j\lambda_j$ is an upper bound on the expected number of undetected defects in the jth subpopulation. If it is the case that once a defect is detected, the afflicted unit is replaced in the population with a healthy unit, then $T^*N_j\lambda_j$ represents a bound on the

expected difference between the number of healthy units and the total number of units in the subpopulation. For λ_j small, as would be the case for a relatively infrequently occurring disorder, this should represent no difficulty. In this paper, therefore, λ_j is considered to be small, which is consistent with the examples of potential applications which are mentioned.

The times of incidence for the defect in the jth subpopulation are designated by the sequence $\{S_j^k\}$ for k = 1, 2,... The screening test is assumed to be administered to the jth subpopulation at the times $1/r_j$, $2/r_j$, $3/r_j$,... The testing frequency r_j is a control variable. The function $p_\ell(\cdot)$ describes the reliability characteristics of the type ℓ screening test, i.e., how likely such a test is to detect a defect as a function of the defect's age. Note that $p_\ell(t) = 0$ for $t < 0$.

The random variable $\overline{S}_{r,\ell}^{k,j}$ denotes the time at which the kth defect is detected. $\overline{S}_{r,\ell}^{k,j}$ depends on the arrival time of the defect (S_j^k), the type of test used (ℓ), and the testing frequency (r_j).

Given the application of test type ℓ at the times $\{1/r_j, 2/r_j,...\}$, the disutility incurred by the kth defect is $D\left(\overline{S}_{r,\ell}^{k,j} - S_j^k\right)$, where $D(\cdot)$ is a strictly increasing function. The total disutility incurred due to those defects which occurred in the interval $[i/r_j, (i+1)/r_j)$, is

$$B_{r,\ell,i} = \sum_{k=1}^{\infty} D\left(\overline{S}_{r,\ell}^{k,j} - S_j^k\right) 1_{i/r_j}(S_j^k) ,$$

where $1_{i/r}(t)$ is an indicator function defined by

$$1_{i/r}(t) = 1_{[i/r,(i+1)/r]}(t) = \begin{cases} 1 & \text{if } i/r \le t \le (i+1)/r \\ 0 & \text{otherwise .} \end{cases}$$

The following theorem, which was proved in the previous paper, provides an expression for $E[B_{r,\ell,i}]$ in terms only of the disutility due to detection delay (as expressed by $D(\cdot)$) and of the reliability of test type ℓ (as expressed by $p_\ell(\cdot)$). In addition, the theorem shows that $E[B_{r,\ell,i}] = E[B_{r,\ell,0}]$ for $i = 0, 1, 2, \ldots$

Theorem A

$$E[B^j_{r,\ell,i}] = N\lambda \sum_{n=1}^{\infty} \int_{\frac{n-1}{r_j}}^{\frac{n}{r_j}} D(u)p_\ell(u) \prod_{m=1}^{n-1} [1 - p_\ell(u - \tfrac{m}{r_j})]\,du$$

for $i = 0, 1, 2, \ldots$

By defining $\bar{B}^j_{r,\ell}$ as the long-run expected disutility per unit time, then by Theorem A

$$\bar{B}^j_{r,\ell} = \lim_{m\to\infty} \frac{r_j}{m} \sum_{i=0}^{m-1} E[B^j_{r,\ell,i}]$$

$$= r_j E[B^j_{r,\ell,0}] = r_j N_j \lambda_j \sum_{n=1}^{\infty} \int_{\frac{n-1}{r_j}}^{\frac{n}{r_j}} D(u)p_\ell(u) \prod_{m=1}^{n-1} [1 - p_\ell(u - \tfrac{m}{r_j})]\,du .$$

(The factor r_j enters the definition to convert disutility per unit testing-interval into per unit time.)

Now it is possible to formulate a mathematical program for the problem of selecting testing frequencies and test-types for each of Q different susceptibility classes which together comprise the whole population. These classes may differ from one another in the number of units they contain, in their defect incidence intensity, and/or in the cost per test application

to an individual for a particular type of test. The subpopulations, however, are assumed to share a common $D(\cdot)$ function. Also it should be noted that the test used may be a function of the jth subpopulation, i.e., $\ell = \ell(j)$.

The multi-subpopulation mass screening problem subject to a budget constraint is:

$$\text{(1)} \qquad \underset{(r_1, r_2, \ldots, r_Q, \ell(1), \ldots, \ell(Q))}{\text{Minimize}} \quad \sum_{j=1}^{Q} \bar{B}^j_{r,\ell(j)}$$

such that

$$\text{(2)} \qquad \sum_{j=1}^{Q} N_j c_{j,\ell(j)} r_j \leq b$$

$$\text{(3)} \qquad r_j > 0 \qquad j = 1, \ldots, Q$$

$$\text{(4)} \qquad \ell(j) \in \mathcal{L} \qquad j = 1, \ldots, Q$$

where

$c_{j,\ell(j)}$ = cost per application of a test of type $\ell(j)$ to an individual of subpopulation j,

N_j = number of units (or individuals) in subpopulation j,

b = budget per unit time.

In order to make this mathematical program even more comprehensive, it is possible to add constraints on the amount of testing labor available and on the capacity of the testing facilities in terms of the number of arrivals, the frequency of testing and the type of tests used. In addition, the cost of false positives can be included as a part of the test costs in inequality (2).

For example, a constraint on the total labor available is:

$$\sum_{j=1}^{Q} N_j \delta_{j,\ell(j)} r_j \leq L_\ell \tag{5}$$

and on the total testing facilities available is:

$$\sum_{j=1}^{Q} N_j f_{j,\ell(j)} r_j \leq F_\ell \tag{6}$$

for each type of test $\ell(\cdot)$ used over the subpopulations being tested, where

$\delta_{j,\ell(j)}$ = amount of labor needed to administer test type $\ell(j)$ to an individual of subpopulation j,

$f_{j,\ell(j)}$ = amount of testing facility time needed to administer test type $\ell(j)$ to an individual of subpopulation j,

L_ℓ = total amount of labor available to administer test type $\ell(\cdot)$ per unit time,

F_ℓ = total amount of facility time available to administer test type $\ell(\cdot)$ per unit time.

This mathematical program may be solved by any of several nonlinear programming methods. However, in certain special, quite important, cases, more explicit formulations of the objective function are possible. The previous paper by Pierskalla and Voelker [1976] extensively considered the case when $p_\ell(t) = 1$ for all $t > 0$. The following sections are concerned with the cases of constant test reliability and with threshold test reliability.

2. Constant Test Reliability

The mass screening model yields interesting results for a test which has a fixed probability p of detecting the disorder if it is present in an individual. Such a model would arise if the unreliability of the test is intrinsic to the test procedure rather than partially dependent upon the state or age of the defect. An example of this is the administration of a Mantoux test for tuberculosis in, say, a population of grade school children. The test has a small but relatively constant level of false negatives. There are other medical tests with similar characteristics (such as electrocardial analysis).

To see how another type of situation with constant test reliability could arise, consider a production process which is subject to a randomly occurring defect which degrades its performance. Once the defect occurs, the level of degradation of the process remains constant, until the defect is discovered. Suppose the defect is such that each item produced has probability δ of being defective, and that the system without the defect never produces defective items.

The only way to discover the existence of the defect in the production system is to examine an item produced which is itself defective. Now, if the examination of an item is expensive (e.g., the item is destroyed as a result of the inspection) and if the capacity to examine a sequence of items involves a setup cost (say a), the following strategy might be called for: At specified times 1/r, 2/r,..., set up the capacity to examine a sequence of items and examine, say, ℓ items at each of those times. The times 1/r, 2/r,... are then the times of testing and the sample size $\ell \in \mathcal{L} = \{1, 2, 3, \ldots\}$ specifies the test type.

Assume that if a defective item is examined, the defect is always observed, and the production process is thereby discovered to be in the degraded state. Hence, a degraded state of the production process will go undetected at the testing occasion k/r if and only if each of the ℓ items sampled at time k/r is, by chance, not defective. But the probability of that event is $(1-\delta)^{\ell}$. Note that the elapsed time t between the entry of the production process into the degraded state and the test time k/r does not affect this probability. Hence,

$$p_{\ell} = p_{\ell}(t) = 1 - (1-\delta)^{\ell}$$

which represents the probability that a test (the inspection of ℓ items) will detect a degraded production process.

If the inspection of a single item has cost c, then to set up and test a sample of ℓ items has cost $c_{\ell} = a + \ell c$. To do this with frequency r, a budget of at least rc_{ℓ} would be required.

The decision maker decides how large a sample of parts to test on each testing occasion, and how frequently to schedule those occasions. If there are a large number of production processes involved, they may be partitioned according to their likelihood of becoming defective--thereby forming susceptibility classes. The decision as to sample size and frequency would then have to be made for each class so as to stay within an overall budget constraint.

This latter problem, several susceptibility classes in a population of production processes, could then be formulated as the program at the end of the previous section. For present purposes, the feature of interest in this example is that $p_{\ell}(t)$ does not depend on t.

Returning to the more general development, recall that $\bar{B}^{j}_{r,\ell}$ represented the long-run expected disutility per unit time if test type ℓ is applied with frequency r_j. Exploiting the fact that $p_\ell(t)$ is a constant (i.e., $p_\ell(t) = p_\ell$ for all $t \geq 0$) and, for ease of analysis, temporarily suppressing the subscript j,

$$\bar{B}_{r,\ell} = N\lambda r p_\ell \sum_{n=0}^{\infty} (1 - p_\ell)^n \int_{\frac{n}{r}}^{\frac{n+1}{r}} D(u)du . \tag{7}$$

In order to avoid confusion with the general case of $\bar{B}_{r,\ell}$ given in the preceding section, it is convenient to use a new notation for (7). Consequently, let p_ℓ be denoted by p and let

$$C(r, p) = N\lambda r p \sum_{n=0}^{\infty} (1 - p)^n \int_{\frac{n}{r}}^{\frac{n+1}{r}} D(u)du . \tag{8}$$

Therefore, (7) and (8) represent the long-run expected disutility per unit time if a test, of a type which is characterized by a reliability which does not vary with time since the defect's incidence, is administered with frequency r.

A question which is examined next is how do changes in r and p affect C(r, p). After that explicit solutions are given for C(r, p) when D(•) takes certain simple forms. And lastly, some general rules are indicated for selecting between a particular kind of test and a more expensive but more reliable alternative test when D(•) takes certain forms.

The following lemma indicates that C(r, p) behaves consistently with one's expectations as r or p varies.

Lemma 1: If $D(\cdot)$ is a strictly increasing function,

$$\frac{\partial C(r, p)}{\partial p} < 0 \quad , \quad \text{and} \tag{9}$$

$$\frac{\partial C(r, p)}{\partial r} < 0 \quad . \tag{10}$$

Note that for $D(\cdot)$ nondecreasing, the inequalities in (9) and (10) still hold but not strictly.

From this lemma, as anticipated, when p increases, the expected disutility decreases. Similarly as r increases, the interval 1/r between tests decreases and the expected disutility decreases. Consequently, as better test types are used or the tests are more frequently applied, the value of such changes in terms of reduced disutility versus the costs of the changes can be assessed and the tradeoffs evaluated.

It is easy to compute the Hessian for $C(r, p)$ when $D(\cdot)$ is differentiable.

$$\frac{\partial^2 C}{\partial r^2} = N\lambda p \sum_{n=0}^{\infty} q^n r^{-3} \left[(n+1)^2 D'\left(\frac{n+1}{r}\right) - n^2 D'\left(\frac{n}{r}\right) \right]$$

$$\frac{\partial^2 C}{\partial p^2} = N\lambda r \sum_{n=2}^{\infty} n(n-1) q^{n-2} [W(n, r) - W(n-1, r)]$$

$$\frac{\partial^2 C}{\partial p \partial r} = N\lambda \sum_{n=0}^{\infty} [1 - (n+1)p] q^{n-1} [W(n, r) - A(n+1, r) + A(n, r)]$$

where $q = 1 - p$; $W(n, r) = \int_{\frac{n}{r}}^{\frac{n+1}{r}} D(s)ds$; $A(n, r) = \frac{n}{r} D\left(\frac{n}{r}\right)$.

Note that if $D'(\cdot)$ is increasing, then $\frac{\partial^2 C}{\partial r^2} \geq 0$ and $\frac{\partial^2 C}{\partial p^2} \geq 0$. Hence, along coordinate directions $C(r, p)$ is convex.

Simple expressions for $C(r; p)$ can be given when $D(\cdot)$ is specialized to a geometric and an exponential function, respectively. Since these two types of functions are reasonably general and are intuitively appealing for disutility functions (in the sense of an increasing rate of disutility the longer the defect is undetected), they can be quite useful as realistic approximations in applications.

Lemma 2: If $D(t) = at^m$ for $a > 0$ and $m \geq 1$, then

$$C(r, p) = \frac{aN\lambda p}{(m+1)r^m} \sum_{n=1}^{\infty} n^{m+1} pq^{n-1} .$$

In the case that m is a positive integer, then the summation in Lemma 2 above is just the moment generating function of a geometric random variable with parameter p and $C(r, p)$ can be expressed by

$$C(r, p) = \frac{aN\lambda p \psi^{(m+1)}(0 : p)}{(m+1)r^m} \tag{11}$$

where

$$\psi(t : p) = \frac{pe^t}{1 - qe^t} \quad \text{and} \quad \psi^{(m)}(0 : p) = \frac{d^m \psi(t : p)}{dt^m}\Big|_{t = 0} .$$

For $m = 1$, $C(r, p) = \frac{aN\lambda}{2r}\left(\frac{2 - p}{p}\right)$, and for $m = 2$, $C(r, p) = \frac{aN\lambda}{3r^2}\left[1 + \frac{6 - 6p}{p^2}\right]$.

Lemma 3: If $D(t) = \beta e^{at}$ for $a, \beta > 0$, then

$$C(r, p) = \beta N\lambda p r\,(e^{a/r} - 1)/a(1 - qe^{a/r})$$

for $r > \frac{-a}{\log(q)}$.

Lemmas 2 and 3 can provide a means to select between two alternative kinds of tests which differ with respect to reliability of detection and cost per application. Let test #1 have cost per application c_1 and reliability p_1. The corresponding parameters for test #2 are c_2 and p_2. If test #1 is administered with frequency rc_2/c_1, and test #2 administered with frequency r, both testing regimes will consume equal quantities of the budgeted resource; viz, Nrc_2. If $C(rc_2/c_1, p_1) \leq C(r, p_2)$ for all $r \geq 0$, then the expected disutility per unit time will be less with test #1 at all levels of budget Nrc_2. That is, if test #2 is being used with frequency r, the expected disutility can be decreased without any additional allotment of budget, simply by switching to test #1 and testing as frequently as the budget permits.

Suppose, for example, that $D(t) = at^m$ for m a positive integer. Then

$$C(rc_2/c_1, p_1) \leq C(r, p_2)$$

if and only if

$$(c_1/c_2)^m \leq \frac{p_2 \psi(0 : p_2)}{p_1 \psi(0 : p_1)} \quad . \tag{12}$$

If the inequality in (12) is reversed, then test #2 is preferred to test #1. Notice that this rule does not depend on r, and hence does not depend on the budget contribution Nrc_2.

3. Threshold Test Reliability

In the previous section the reliability of the test depended only on factors intrinsic to the test itself and did not depend at all on the elapsed time since incidence at the time of the test. In this section a special form for $p(t)$ is considered which is very different from the case of constant test reliability. Here the test reliability is zero, if the elapsed time since the defect's incidence is less than T, and that at elapsed time T, the reliability jumps to one. That is, $p(t) = 1_{[T,\infty)}(t)$ where the number T is a characteristic of the type of test chosen.

The primary results in this section are a simple characterization of $\overline{B}_{r,\ell}$ when $p_\ell(\cdot) = 1_{[T_\ell,\infty)}(\cdot)$, and rules which, in some cases, will permit selection between two tests which differ in their reliability (i.e., in their detection threshold T) and in their cost per application.

Since the number T_ℓ completely specifies the function $p_\ell(t) = 1_{[T_\ell,\infty)}(t)$, the long-run expected disutility per unit time is expressed in terms of r and T rather than in terms of r and the test type index ℓ. Suppressing the subpopulation index j and the test index ℓ, define

$$A(r, T) = \overline{B}_{r,\ell} \text{ when } p_\ell(\cdot) = 1_{[T_\ell,\infty)}(\cdot) . \tag{13}$$

When $T = 0$, this definition is equivalent to perfect test reliability or $p_\ell(t) = 1$ for $t > 0$ and

$$A(r, 0) = rN\lambda \int_0^{\frac{1}{r}} D(u)du .$$

The next theorem evaluates $A(r, T)$ for all values of $T \geq 0$. Of course, the "blind period" of the test from 0 to T does not delay the

detection of each arriving defect exactly T units of time. The amount of delay depends on the interplay among the time of arrival of the defect, the testing schedule $\{1/r, 2/r, \ldots\}$, and the magnitude of T.

Theorem 1: If $p(\cdot) = 1_{[T,\infty)}(\cdot)$ for some $T \geq 0$, then

$$A(r, T) = N\lambda r \int_0^{\frac{1}{r}} D(T+u)du \,. \tag{14}$$

Suppose the decision maker has two kinds of tests available, and he must choose one of them for implementation in a mass screening program. Suppose the first kind of test--call it test #1--has sensitivity characterized by the "time-until-detectability" threshold T_1. That is, this test will detect a disorder if and only if the disorder has been present for a length of time T_1. Let $c_1 > 0$ be the cost per application (to an individual unit) of this kind of test. For the second kind of test under consideration, test #2, let T_2 and c_2 be the corresponding parameters.

Assume test #1 is better in the sense that $T_1 < T_2$. If $c_1 \leq c_2$, there is no decision to make; test #1 should be used. Therefore, assume $c_1 > c_2$.

If the exact shape of the function $D(\cdot)$ is known, Theorem 1 can be used to decide which test to use for each possible level of budget. Let b be the budget per individual in the (homogeneous) population. Then the use of test #1 will permit a testing frequency of b/c_1, and the use of test #2 permits frequency b/c_2. To decide which test to use, compare the expected disutilities per unit time assuming a fully allocated budget, i.e., compare $A(b/c_1, T_1)$ and $A(b/c_2, T_2)$. With $D(\cdot)$ known, these quantities can be evaluated explicitly by Theorem 1 and compared.

Note that it was assumed that the entire budget should be allocated. This assumption was unnecessary because when $D(\cdot)$ is an increasing function

$$\frac{\partial}{\partial r} A(r, p) = N\lambda \left[\int_0^{\frac{1}{r}} D(T+u) - D(T+\tfrac{1}{r}) du \right] < 0 .$$

Therefore, the greater the testing frequency, the lower the expected disutility per unit time. For this reason, the budget should be entirely allocated.

When the exact form of the disutility function is not known, Theorem 1 does not directly differentiate between test #1 and test #2. However, the two following theorems will permit such a determination, at least for certain relative configurations of the budget (per unit population), and the relative sensitivity $T_1 - T_2$ and cost differential $c_2 - c_1$ of the tests.

Specifically, Theorem 2 will show that for any (increasing) disutility function test #1 is indicated, if the budget (per unit population) exceeds $(c_1 - c_2)/(T_2 - T_1)$. On the other hand, Theorem 3 shows that for a convex increasing disutility function, test #2 is better if the budget is less than $\frac{c_1 - c_2}{2(T_2 - T_1)}$.

A decision rule for the case where the budget falls between $\frac{c_1 - c_2}{T_2 - T_1}$ and $\frac{c_1 - c_2}{2(T_2 - T_1)}$ has not been found for general disutility functions.

Just how the statements of Theorems 2 and 3 are translated into the above decision rules is explained after the statements of the respective theorems.

Theorem 2: Given $D(\cdot)$ a strictly increasing function, $T_1 < T_2$ and $c_1 > c_2$, then $T_2 - T_1 > (c_1 - c_2)/rc_2$ provides a sufficient condition for

$$A(rc_2/c_1, T_1) < A(r, T_2) \tag{15}$$

and hence for test #1 to be preferable at the per unit population budget level of rc_2.

The hypothesis of Theorem 2 is equivalent to

$$r > \frac{c_1 - c_2}{c_2(T_2 - T_1)} \quad . \tag{16}$$

If test #1 is used with frequency rc_2/c_1 (as on the left side of (15)), the budget required would be rc_2 which is the same budget required to execute test #2 with frequency r (as on the right side of (15)).

Hence, translating (16) into its equivalent form in terms of budget (per unit population), designated by b,

$$b = rc_2 > (c_1 - c_2)/(T_2 - T_1) \, .$$

With its hypothesis in this form, the theorem provides a lower bound on the budget which is a sufficient condition for test #1 to entail lower expected disutility per unit time vis-a-vis test #2, were the two tests scheduled at their respective maximal frequencies rc_2/c_1 and r consistent with the budget rc_2. The following lemma is needed for the proof of Theorem 3 and is recorded here for general interest.

Lemma 4: If f is a convex function, then

$$\frac{1}{y}\int_t^{t+y} f(s)ds \leq \frac{1}{2}[f(t) + f(t+y)] \, . \tag{17}$$

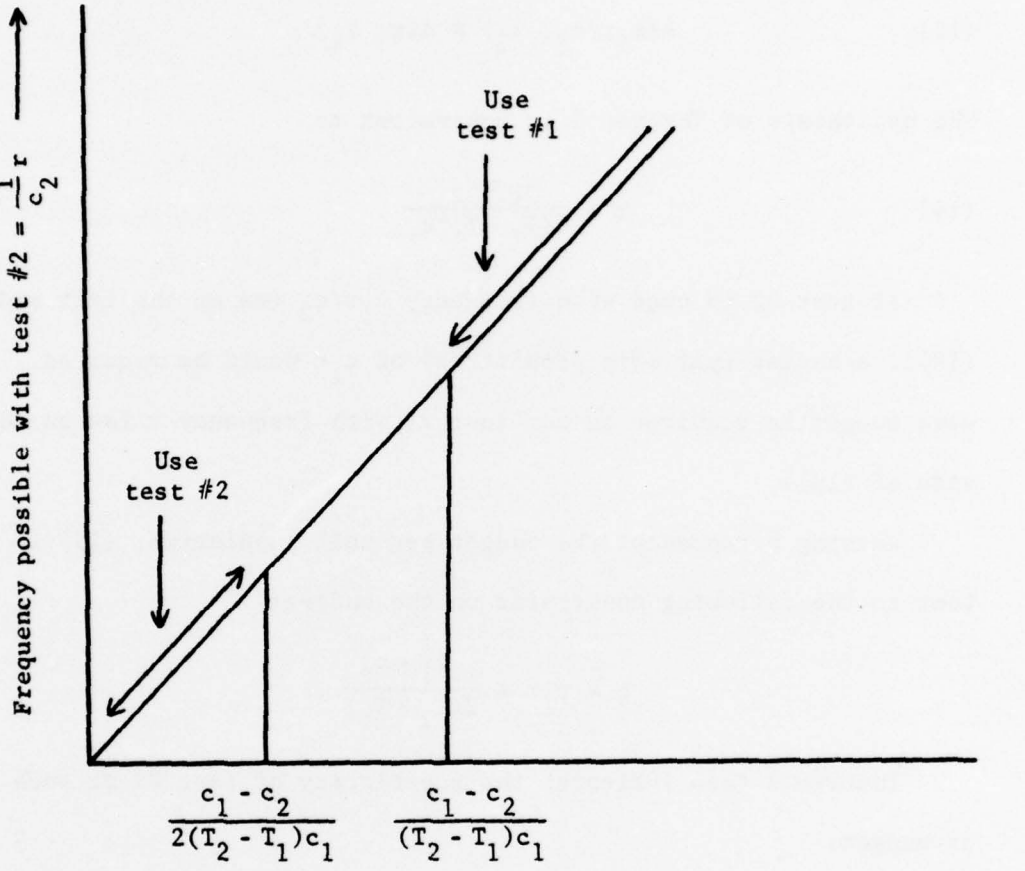


Figure 1

This Figure illustrates Theorems 2 and 3. Each point on the sloping line corresponds to a level of budget. Regions where each test is preferable are indicated.

Theorem 3: If $D(\cdot)$ is convex and increasing, $c_1 > c_2$, $T_1 < T_2$, and $T_2 - T_1 \leq (c_1 - c_2)/2rc_1$, then

$$A(c_1 r/c_2,\ T_2) \leq A(r,\ T_1)\ . \tag{18}$$

The hypothesis of Theorem 3 is equivalent to

$$r \leq \frac{c_1 - c_2}{2(T_2 - T_1)c_1}\ . \tag{19}$$

If test #2 is used with frequency $c_1 r/c_2$ (as on the left side of (18)), a budget (per unit population) of $c_1 r$ would be required. The same budget is required to use test #1 with frequency r (as on the right side of (18)).

Letting b represent the budget per unit population, (19) is equivalent to the following constraint on the budget:

$$b = c_1 r \leq \frac{c_1 - c_2}{2(T_2 - T_1)}\ .$$

Theorem 3 then indicates the superiority of test #2 at such levels of budget.

4. Some Examples of Disutility Functions

Suppose a production process is subject to a randomly occurring defect. Although production appears to proceed normally after the incidence of the defect, the product produced is thereafter defective to an extent which remains constant until the production process is returned to its proper mode of operation. The only way to learn if the production process is in this degraded state is to perform a costly test. Now, if a test detects the existence of the degraded mode of production t units of time

after its incidence, the harm done will be proportional to the amount of defective product (unknowingly) produced which, in turn, is proportional to t. Hence $D(t) = at$ for some $a > 0$.

Another example where a linear $D(\cdot)$ function may be appropriate would be for the periodic inspection of an inactive device (such as a missile) stored for possible use in an emergency. If t is the time between the incidence of the disorder and its detection, the disutility incurred may be proportional to the probability that the device would be needed in that time interval. If such "emergencies" arise according to a Poisson process with rate μ, then the probability of an emergency in a time interval of length t is $1 - e^{-\mu t}$, which for μ small is approximately μt. Hence, if b is the cost incurred should there be an emergency while the device is defective, and if μ is the (small) arrival rate of emergencies, then $D(t) = b\mu t$.

A quadratic disutility could arise in the following situation. Suppose the magnitude of a randomly occurring defect increases linearly with time since the occurrence of the defect. For example, the magnitude of the defect might be the size of a small leak in a storage container for a fluid, and as fluid escapes the leak gets larger. Further suppose that the harm done accumulates at a rate proportional to the magnitude of the defect. Hence, the quantity of fluid lost (at least initially) increases the longer the defect exists, and the rate of fluid loss is proportional to the size of the leak.

Let the size of the leak (as measured by rate of fluid loss), at time s since the leak's incidence, be cs. Then, if the defect is detected at time t since incidence, the disutility incurred (fluid lost) is

$$D(t) = \int_0^t csds = \frac{1}{2}ct^2 .$$

Examples of nonlinear disutility functions for a few animal and human diseases are already known. Although data may not be available currently to support the utilization of particular models, that should not deter the development of such models. It is reasonable to expect that in many application areas in the future much additional data will become available; for example, data concerning the stochastic pattern of a particular disease's development. Furthermore, these models can serve as a guide to the kinds of data which should be gathered.

It is also reasonable to expect the future development of improved testing technologies capable of detecting a disorder in a much earlier stage of development than is now the case. Such innovations will make screening programs more attractive and, consequently, make more important the analytical tools to design such programs intelligently.

Appendix

Proof of Lemma 1: Letting $q = 1 - p$,

$$\frac{\partial}{\partial p} C(n, p) = N\lambda r \sum_{n=0}^{\infty} q^n W(n) - N\lambda rp \sum_{n=0}^{\infty} nq^{n-1} W(n)$$

where $$W(n) = \int_{\frac{n}{r}}^{\frac{n+1}{r}} D(s)ds .$$

Now let

$$V_0 = W_0$$

$$V_n = W_n - W_{n-1}, \quad n = 1, 2, \ldots$$

Then, $W_n = \sum_{j=0}^{n} V_j$ and

$$\frac{\partial}{\partial p} C(r, p) = N\lambda r \sum_{n=0}^{\infty} q^n \sum_{j=0}^{n} V_j - N\lambda rp \sum_{n=0}^{\infty} nq^{n-1} \sum_{j=0}^{n} V_j$$

$$= N\lambda r \sum_{j=0}^{\infty} V_j \left[\sum_{n=j}^{\infty} q^n - p \sum_{n=j}^{\infty} nq^{n-1} \right]$$

$$= -N\lambda r \sum_{j=0}^{\infty} jq^{j-1} V_j$$

$$= N\lambda r \sum_{j=1}^{\infty} jq^{j-1} \left[\int_{\frac{j-1}{r}}^{\frac{j}{r}} D(s)ds - \int_{\frac{j}{r}}^{\frac{j+1}{r}} D(s)ds \right] , \tag{A1}$$

since

$$\sum_{n=j}^{\infty} q^n - p \sum_{n=j}^{\infty} nq^{n-1} = q^{j-1}[q/p - j - q/p] = -jq^{j-1} .$$

Because $D(\cdot)$ is strictly increasing, the right-hand side of (A1) is negative, and $\frac{\partial}{\partial p} C(r, p) < 0$.

$$\frac{\partial}{\partial r} C(r, p) = N\lambda p \sum_{n=0}^{\infty} q^n \int_{\frac{n}{r}}^{\frac{n+1}{r}} D(s)ds$$

$$- N\lambda p \sum_{n=0}^{\infty} q^n \left(\frac{n+1}{r} D\left(\frac{n+1}{r}\right) - \frac{n}{r} D\left(\frac{n}{r}\right) \right)$$

$$= N\lambda p \sum_{n=0}^{\infty} q^n \left[\int_{\frac{n}{r}}^{\frac{n+1}{r}} D(s)ds - \frac{n}{r}\left(D\left(\frac{n+1}{r}\right) - D\left(\frac{n}{r}\right)\right) - \frac{1}{r} D\left(\frac{n+1}{r}\right)\right] < 0 .$$

This inequality follows from $D(\cdot)$ increasing, through the relations

$$D\left(\frac{n+1}{r}\right) - D\left(\frac{n}{r}\right) > 0$$

and

$$\int_{\frac{n}{r}}^{\frac{n+1}{r}} D(s)ds - \frac{1}{r} D\left(\frac{n+1}{r}\right) < \int_{\frac{n}{r}}^{\frac{n+1}{r}} D\left(\frac{n+1}{r}\right)ds - \frac{1}{r} D\left(\frac{n+1}{r}\right) = 0 .$$

Q.E.D.

Proof of Lemma 2:

$$C(r, p) = aN\lambda rp \sum_{n=0}^{\infty} q^n \int_{\frac{n}{r}}^{\frac{n+1}{r}} t^m dt$$

$$= \frac{aN\lambda p}{(m+1)r^m} \left[\sum_{n=0}^{\infty} q^n (n+1)^{m+1} - \sum_{n=0}^{\infty} q^n n^{m+1} \right]$$

$$= \frac{aN\lambda p}{(m+1)r^m} \left[\sum_{n=1}^{\infty} q^{n-1} n^{m+1} - \sum_{n=1}^{\infty} q^n n^{m+1} \right]$$

$$= \frac{aN\lambda p}{(m+1)r^m} \sum_{n=1}^{\infty} n^{m+1} pq^{n-1} .$$

Q.E.D.

Proof of Lemma 3:

$$C(r, p) = N\lambda p \sum_{k=0}^{\infty} q^k \frac{r}{a} e^{ak/r}(e^{a/r} - 1)$$

$$= \frac{N\lambda pr}{a}(e^{a/r} - 1) \sum_{k=0}^{\infty} (qe^{a/r})^k .$$

The geometric series $\sum_{k=0}^{\infty} (qe^{a/r})^k$ converges if and only if $qe^{a/r} < 1$, i.e., if and only if

$$r > \frac{-a}{\log q} .$$

Therefore, for $r > \frac{-a}{\log q}$

$$C(r, p) = \frac{N\lambda pr}{a}(e^{a/r} - 1) \frac{1}{1 - qe^{a/r}} .$$

Q.E.D.

Proof of Theorem 1: The following conventions are used below:

$\prod_{i=1}^{n} x_i = 1$ for $n < 1$ and $[x]$ is the largest integer not exceeding x.

From the definition of $\bar{B}_{r,\ell}$,

$$\text{(A2)} \qquad A(r, T) = N\lambda r \sum_{n=1}^{\infty} \int_{\frac{n-1}{r}}^{\frac{n}{r}} D(u) 1_{[T,\infty)}(u) \prod_{m=1}^{n-1} \left[1 - 1_{[T,\infty)}(u - \tfrac{m}{r})\right] du .$$

Let

$$\text{(A3)} \qquad \Lambda(u) = \prod_{m=1}^{[ru]} \left[1 - 1_{[T,\infty)}(u - \tfrac{m}{r})\right]$$

Note that for $u \in [(n-1)/r, n/r)$, $n - 1 = [ru]$.

Therefore, (A2) becomes

$$A(r, T) = N\lambda r \sum_{n=1}^{\infty} \int_{\frac{n-1}{r}}^{\frac{n}{r}} D(u) 1_{[T,\infty)}(u) \Lambda(u) du . \tag{A4}$$

Now it will be shown that

$$\Lambda = 1_{[0, T+\frac{1}{r})} . \tag{A5}$$

Select $x \geq T+\frac{1}{r}$. Then $[rx] \geq 1$ and by (A3)

$$\Lambda(x) = \prod_{m=1}^{[rx]} \left[1 - 1_{[T,\infty)}(x - \tfrac{m}{r}) \right] \leq 1 - 1_{[T,\infty)}(x - \tfrac{1}{r}) .$$

But $x - \frac{1}{r} \geq T$, therefore the above relation shows

$$\Lambda(x) = 0 \quad \text{for } x \geq T+\tfrac{1}{r} . \tag{A6}$$

Now select $y \in [0, T+\frac{1}{r})$. If $[ry] = 0$, (A3) and the convention on degenerate products gives $\Lambda(y) = 1$. On the other hand, if $[ry] > 0$, then the relation $y - \frac{m}{r} \leq y - \frac{1}{r} < T$ for $m = 1, 2, \ldots, [ry]$ is not vacuous (i.e., the set of allowable indices m is not empty) and neither is the consequent relation

$$1_{[T,\infty)}(y - \tfrac{m}{r}) = 0 \quad \text{for } m = 1, 2, \ldots, [ry] .$$

And, hence by (A3)

$$\Lambda(y) = 1 \quad \text{for } y \in [0, T+\tfrac{1}{r}) . \tag{A7}$$

(A5) now follows from (A6) and (A7). Substituting (A5) into (A4),

$$A(r, T) = N\lambda r \sum_{n=1}^{\infty} \int_{\frac{n-1}{r}}^{\frac{n}{r}} D(u) 1_{[T,\infty)}(u) 1_{[0,T+\frac{1}{r})}(u) du$$

$$= N\lambda r \int_0^{\infty} D(u) 1_{[T,T+\frac{1}{r})}(u) du$$

$$= N\lambda r \int_0^{\frac{1}{r}} D(T+s) ds .$$

Q.E.D.

Proof of Theorem 2:

Suppose $t \to D(t)$ is the underlying disutility function. Then, by Theorem 1, (15) is equivalent to showing

$$\text{(A8)} \qquad \frac{rc_2}{c_1} \int_0^{c_1/rc_2} D(T_1 + s) ds < r \int_0^{\frac{1}{r}} D\Big(T_1 + (T_2 - T_1) + s\Big) ds .$$

Let

$$D^*(t) = D(T_1 + t) \quad \text{and} \quad T = T_2 - T_1 ,$$

then (A8) becomes

$$\text{(A9)} \qquad \frac{rc_2}{c_1} \int_0^{c_1/rc_2} D^*(s) ds < r \int_0^{\frac{1}{r}} D^*(T+s) ds .$$

Since $D(\cdot)$ is an arbitrary increasing function, so is $D^*(\cdot)$; consequently, the asterisk on $D^*(\cdot)$ may be dropped.

To prove the theorem, it suffices to verify (A9). By hypothesis,

$$T > (c_1 - c_2)/rc_2$$

which implies

$$T > u(c_1 - c_2)/c_2 \qquad \text{for } u \in [0, \tfrac{1}{r}]$$

which implies

$$\text{(A10)} \qquad T+u > uc_1/c_2 \qquad \text{for } u \in [0, \tfrac{1}{r}] .$$

Now,

$$\frac{rc_2}{c_1}\int_0^{c_1/rc_2} D(s)ds = r\int_0^{\frac{1}{r}} D(c_1 s/c_2)ds < r\int_0^{\frac{1}{r}} D(T+s)ds .$$

The inequality follows from (A10) and the fact that $D(\cdot)$ is increasing.

Q.E.D.

Proof of Lemma 4:

Let

$$h(t+s) = f(t) + \frac{s}{y}[f(t+y) - f(t)] \quad , \quad s \in [0, y] .$$

Since f is convex,

$$f(t+s) \le h(t+s) , \quad s \in [0, y] .$$

Hence,

$$\frac{1}{y}\int_t^{t+y} f(s)ds \le \frac{1}{y}\int_t^{t+y} h(s)ds = \frac{1}{y}\int_0^{y} h(t+s)ds$$

$$= \frac{1}{y}\int_0^{y} \left[f(t) + \frac{s}{y}\Big(f(t+y) - f(t)\Big)\right] ds$$

$$= f(t) + \frac{1}{y^2}[f(t+y) - f(t)](y^2/2)$$

$$= \frac{1}{2}[f(t+y) + f(t)] .$$

Q.E.D.

Proof of Theorem 3:

Theorem 1 renders the inequality (18) equivalent to

$$\text{(A11)} \qquad \frac{c_1 r}{c_2} \int_0^{c_2/c_1 r} D(T_2 + s)ds \leq r \int_0^{\frac{1}{r}} D(T_1 + s)ds .$$

As in the proof of Theorem 3.6, let

$$D^*(t) = D(T_1 + t) \quad \text{and} \quad T = T_2 - T_1 .$$

(A11) then becomes

$$\text{(A12)} \qquad \frac{c_1 r}{c_2} \int_0^{c_2/c_1 r} D^*(T + s)ds \leq r \int_0^{\frac{1}{r}} D^*(s)ds .$$

Note that $D(\cdot)$ being an arbitrary convex increasing function implies $D^*(\cdot)$ is an arbitrary function of the same class. Hence, the asterisk may be dropped. Also, let $a = c_2/c_1$ and $x = \frac{1}{r}$. Then, (A12) can be written

$$\text{(A13)} \qquad \frac{1}{ax} \int_0^{ax} D(T + s)ds \leq \frac{1}{x} \int_0^{x} D(s)ds .$$

Using Lemma 4

$$\begin{aligned} \frac{1}{ax} \int_0^{ax} D(T + s)ds &= \frac{1}{ax} \int_T^{T+ax} D(s)ds \\ &\leq (1/2)[D(T) + D(T + ax)] \\ &= (1/2T) \int_0^{T} [D(T) + D(T + ax)]ds \\ &= (1/2T) \int_0^{T} G(0)ds \end{aligned} \qquad \text{(A14)}$$

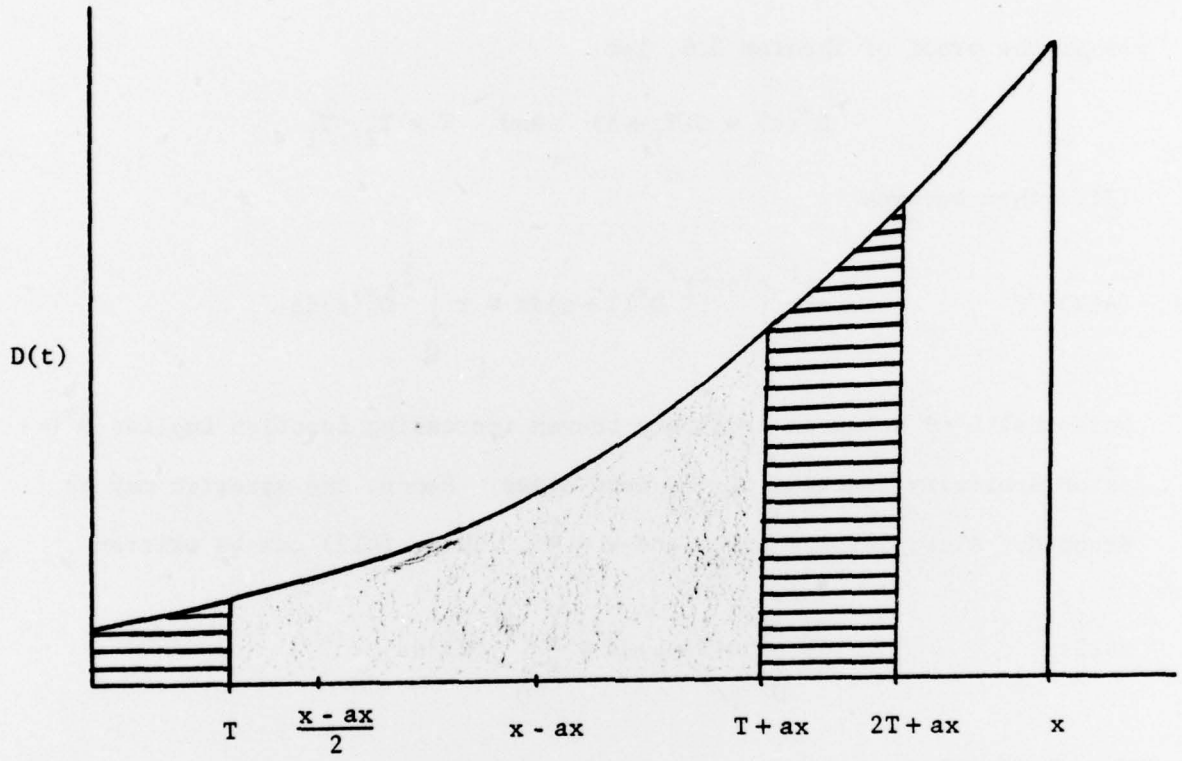


Figure 3.1

This Figure illustrates the regions of integration used in the proof of Theorem 3.

where $G(s) = D(T-s) + D(T+ax+s)$. Now $D(\cdot)$ convex and increasing implies $G(\cdot)$ is convex and increasing in s. Therefore, when $s \geq 0$, $G(0) \leq G(s)$ which, along with (A14), implies

$$\frac{1}{ax}\int_0^{ax} D(T+s)ds \leq \frac{1}{2T}\int_0^{T} G(0)ds \leq \frac{1}{2T}\int_0^{T} G(s)ds$$

$$= \frac{1}{2T}\int_0^{T} [D(T-s) + D(T+ax+s)]ds \,. \tag{A15}$$

Now,

$$\frac{1}{2T+ax}\int_0^{2T+ax} D(s)ds = \frac{1}{2T+ax}\left[\int_T^{T+ax} D(s)ds + \int_0^{T} D(s)ds + \int_{T+ax}^{2T+ax} D(s)ds\right]$$

$$= \frac{1}{2T+ax}\left[\int_T^{T+ax} D(s)ds + \int_0^{T} D(T-s)ds + \int_0^{T} D(T+ax+s)ds\right]$$

$$= \frac{1}{2T+ax}\int_T^{T+ax} D(s)ds + \frac{1}{2T+ax}\int_0^{T} [D(T-s) + D(T+ax+s)]ds$$

$$= \left(\frac{ax}{2T+ax}\right)\frac{1}{ax}\int_T^{T+ax} D(s)ds + \left(\frac{2T}{2T+ax}\right)\frac{1}{2T}\int_0^{T} [D(T-s) + D(T+ax+s)]ds$$

$$\geq \frac{1}{ax}\int_T^{T+ax} D(s)ds \,. \tag{A16}$$

This last inequality follows from

$$\frac{1}{ax}\int_{T}^{T+ax} D(s)ds \leq \frac{1}{2T}\int_{0}^{T} [D(T-s) + D(T+ax+s)]ds$$

which follows from (A15). By hypothesis, $T \leq x(c_1 - c_2)/2c_1$ which implies

(A17) $$2T + ax \leq x .$$

$$\frac{1}{x}\int_{0}^{x} D(s)ds = \frac{1}{x}\int_{0}^{2T+ax} D(s)ds + \frac{1}{x}\int_{2T+ax}^{x} D(s)ds$$

$$= \left(\frac{2T+ax}{x}\right)\frac{1}{2T+ax}\int_{0}^{2T+ax} D(s)ds$$

$$+ \left(\frac{x-2T-ax}{x}\right)\frac{1}{x-2T-ax}\int_{2T+ax}^{x} D(s)ds$$

(A18) $$\geq \frac{1}{2T+ax}\int_{0}^{2T+ax} D(s)ds .$$

The last inequality follows from

$$\frac{1}{2T+ax}\int_{0}^{2T+ax} D(s)ds \leq \frac{1}{x-2T-ax}\int_{2T+ax}^{x} D(s)ds$$

which follows from (A17) since $D(\cdot)$ is increasing. The proof is completed by using, in turn, (A16) and (A18) to obtain

$$\frac{1}{ax}\int_{0}^{ax} D(T+s)ds = \frac{1}{ax}\int_{T}^{T+ax} D(s)ds$$

$$\leq \frac{1}{2T+ax}\int_{0}^{2T+ax} D(s)ds \leq \frac{1}{x}\int_{0}^{x} D(s)ds .$$

Q.E.D.

REFERENCES

Barlow, R., L. Hunter, and F. Proschan, "Optimum Checking Procedures," J. Soc. Ind. Appl. Math., 11, 4 (1963), pp. 1078-1095.

Blumberg, M. S., "Evaluating Health Screening Procedures," Operations Research, 5 (1957), p. 351.

Çinlar, E., Introduction to Stochastic Processes, Englewood Cliffs, New Jersey: Prentice-Hall, 1975.

Derman, C., "On Minimax Surveillance Schedules," Naval Research Logistics Quarterly, 8 (1961), pp. 415-419.

Flagle, C. D., "A Decision Theoretical Comparison of Three Procedures of Screening for a Single Disease," Proceedings of the Fifth Berkeley Symposium (1967), p. 887.

Hutchison, G. and S. Shapiro, "Lead Time Gained by Diagnostic Screening for Breast Cancer," Journal of the National Cancer Institute, 41 (1968), p. 665.

Keller, J., "Optimum Checking Schedules for Systems Subject to Random Failure," Management Science, 21 (1974), pp. 256, 260.

Kirch, R. and M. Klein, "Surveillance Schedules for Medical Examinations," Management Science, 20 (1974), pp. 1403, 1409.

_____, "Examination Schedules for Breast Cancer," Cancer, 33 (1974), pp. 1444, 1450.

Lincoln, T. and G. H. Weiss, "A Statistical Evaluation of Recurrent Medical Examinations," Operations Research, 12 (1964), pp. 187, 205.

McCall, J., "Maintenance Policies for Stochastically Failing Equipment: A Survey," Management Science, 11 (1965), pp. 493-524.

_____, "Preventive Medicine Policies," Rand Corporation, P-3368-1, 1969.

Neveu, J., Mathematical Foundations of the Calculus of Probability, San Francisco: Holden-Day, 1965.

Pierskalla, W. P. and J. A. Voelker, "A Survey of Maintenance Models: The Control and Surveillance of Deteriorating Systems," Naval Research Logistics Quarterly, 23, 3 (1976), pp. 353-388.

_____, "A Model for Optimal Mass Screening and the Case of Perfect Test Reliability," Technical Report No. 3, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois 60201, 1977.

Prorok, P., "On the Theory of Periodic Screening for the Early Detection of Disease," Unpublished Ph.D. dissertation, University of New York at Buffalo, 1973.

_____, "The Theory of Periodic Screening, I: Lead Time and Proportion Detected," Advances in Applied Probability, 8 (March, 1976), pp. 127-143.

_____, "The Theory of Periodic Screening, II: Doubly Bounded Recurrence Times and Mean Lead Time and Detection Probability Estimation," Advances in Applied Probability, 8 (September, 1976), pp. 460-476.

Roeloffs, R., "Minimax Surveillance Schedules with Partial Information," Naval Research Logistics Quarterly, 10 (1963), pp. 307-322.

_____, "Minimax Surveillance Schedules for Replacement Units," Naval Research Logistics Quarterly, 14 (1967), pp. 461-471.

Shwartz, M. and H. Galliher, "Analysis of Serial Screening in an Asymptomatic Individual to Detect Breast Cancer," Technical Report, Department of Industrial and Operations Engineering, College of Engineering, University of Michigan, 1975.

Voelker, J. A., "Contributions to the Theory of Mass Screening," Ph.D. dissertation, Northwestern University, 1976.

Weiss, G. and M. Zelen, "A Semi-Markov Model for Clinical Trials," Journal of Applied Probability, 2 (1965), pp. 269-285.

Zelen, M. and M. Feinleib, "On the Theory of Screening for Chronic Diseases," Biometrika, 56 (1969), p. 601.

Zelen, M., "Problems in the Early Detection of Disease and the Finding of Fault," presented at the meeting of the International Statistics Institute, Washington, D. C., 1971.

20.

cont →

In both cases, the defect arrival process is Poisson.

In the first case, the reliability of the test, i.e., the probability that the test will detect the disease, is dependent only on the type of test being administered and is constant over all values of elapsed time since the incidence of the defect. The long-run expected disutility per unit time is derived and variations of the disutility with regard to test reliability and testing frequency are presented. In addition, explicit solutions are provided for various special forms of the disutility function.

In the second case, the test will detect the defect if and only if the elapsed time since incidence exceeds a critical threshold T which characterizes the test. An expression for the long-run expected disutility per unit time is derived and decision rules to select between two alternative test types and costs are presented.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report Number 5	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Optimal Mass Screening under Constant and Threshold Test Reliabilities		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William P. Pierskalla John A. Voelker		8. CONTRACT OR GRANT NUMBER(s) N00014-67-A-0356-0030 N00014-75-C-0797
9. PERFORMING ORGANIZATION NAME AND ADDRESS Dept. of Industrial Engineering and Management Sciences Northwestern University Evanston, Illinois 60201		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-322
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics and Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE March 31, 1977 13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Maintainability, Inspection, Mass Screening, Preventive Maintenance, Surveillance of Equipment

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper considers a population (composed, for example, of people, machines, or livestock) subject to a randomly occurring defect or disease. If there exist testing procedures capable of detecting the defect before it would otherwise become known, and if such early detection provides benefit, the periodic administration of such a test procedure to the members of the population, i.e., a mass screening program, may be advisable. This paper develops and analyzes two important cases of a general model of a mass screening program. → next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)


FILM